

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| Result No. | Score | Query  | Match | Length | DB       | ID        | Description |            |
|------------|-------|--------|-------|--------|----------|-----------|-------------|------------|
| 1          | 1470  | 100..0 | 1470  | 6      | AX154641 | AX154641  | Sequence    |            |
| 2          | 505   | 34..4  | 2188  | 8      | CAU64206 | U64206    | Candida alb |            |
| 3          | 155   | 10..4  | 2188  | 8      | AF001978 | AF001978  | Candida a   |            |
| c          | 4     | 136..2 | 9..3  | 42565  | 8        | CAC35A5   | AL033396    | C. albican |
| c          | 5     | 91..2  | 6..2  | 86872  | 3        | PPMAL3P5  | AL034556    | Plasmoidiu |
| c          | 6     | 86..8  | 5..9  | 104992 | 2        | AC005504  | AC005504    | Plasmoidiu |
| c          | 7     | 86..8  | 5..9  | 169546 | 2        | AC004157  | AC004157    | Plasmoidiu |
| c          | 8     | 75..2  | 5..1  | 104014 | 2        | AC116921  | AC116921    | Dictyoste  |
| c          | 9     | 73..4  | 5..0  | 14876  | 3        | AE001398  | AE001398    | Plasmoidiu |
| c          | 10    | 73..4  | 5..0  | 163399 | 3        | PPMAL3P6  | 298551      | Plasmoidiu |
| c          | 11    | 72..8  | 5..0  | 11880  | 3        | PPMAL3P4  | AL008970    | Plasmoidiu |
| c          | 12    | 70..8  | 4..8  | 2009   | 6        | AX457067  | AX457067    | Sequence   |
| c          | 13    | 70..4  | 4..8  | 192929 | 2        | AC005505  | AC005505    | Plasmoidiu |
| c          | 14    | 69..4  | 4..7  | 56099  | 2        | AC11598   | AC11598     | Dictyoste  |
| c          | 15    | 69..2  | 4..7  | 25305  | 3        | PPMAL3P7  | AL034559    | Plasmoidiu |
| c          | 16    | 68..2  | 4..6  | 11589  | 2        | AC117072  | AC117072    | Dictyoste  |
| c          | 17    | 68..2  | 4..6  | 9810   | 6        | AX345328  | AX345328    | Sequence   |
| c          | 18    | 67..2  | 4..6  | 7347   | 1        | AF211124  | AF211124    | Carsonell  |
| c          | 19    | 67..2  | 4..6  | 67570  | 3        | PPMAL1P3  | AL031746    | Plasmoidiu |
| c          | 20    | 67..2  | 4..6  | 130593 | 2        | AC001593  | AC001593    | Homo sapi  |
| c          | 21    | 67..2  | 4..6  | 198431 | 9        | AC109994  | AC109994    | Homo sapi  |
| c          | 22    | 66..8  | 4..5  | 917371 | 8        | ATV28AB   | AL162651    | Arabidops  |
| c          | 23    | 66..4  | 4..5  | 1141   | 6        | AX083744  | AX083744    | Sequence   |
| c          | 24    | 66..4  | 4..5  | 14635  | 3        | AE001423  | AE001423    | Plasmoidiu |
| c          | 25    | 66..4  | 4..5  | 18624  | 6        | AX346604  | AX346604    | Sequence   |
| c          | 26    | 66..4  | 4..5  | 30726  | 1        | AT217269  | AT217269    | Dictyoste  |
| c          | 27    | 66..2  | 4..5  | 1192   | 9        | HS4323759 | AL1323759   | Homo sapi  |
| c          | 28    | 66..2  | 4..5  | 14867  | 3        | AE001398  | AE001398    | Plasmoidiu |
| c          | 29    | 66..2  | 4..5  | 11589  | 2        | AC117072  | AC117072    | Dictyoste  |
| c          | 30    | 66..2  | 4..5  | 133148 | 9        | AC009435  | AC009435    | Homo sapi  |
| c          | 31    | 66..6  | 4..5  | 157538 | 6        | AX346059  | AX346059    | Sequence   |
| c          | 32    | 66..4  | 4..5  | 201219 | 2        | AC117073  | AC007158    | Homo sapi  |
| c          | 33    | 65..8  | 4..5  | 156609 | 2        | AC004153  | AC004153    | Plasmoidiu |
| c          | 34    | 65..6  | 4..5  | 7584   | 6        | AX286935  | AX286935    | Sequence   |
| c          | 35    | 65..4  | 4..4  | 22977  | 9        | AC106883  | AC106883    | Homo sapi  |
| c          | 36    | 65..4  | 4..4  | 162515 | 9        | AC107032  | AC107032    | Homo sapi  |
| c          | 37    | 65..2  | 4..4  | 23576  | 2        | AC115599  | AC115599    | Dictyoste  |
| c          | 38    | 65..2  | 4..4  | 124820 | 2        | AC117073  | AC117073    | Dictyoste  |
| c          | 39    | 65..2  | 4..4  | 141017 | 2        | AC116962  | AC116962    | Dictyoste  |
| c          | 40    | 65..2  | 4..4  | 165260 | 9        | AC0424341 | AC024341    | Homo sapi  |
| c          | 41    | 65..2  | 4..4  | 43939  | 2        | AC116965  | AC116965    | Dictyoste  |
| c          | 42    | 65..2  | 4..4  | 165059 | 9        | AC116985  | AC116985    | Homo sapi  |
| c          | 43    | 65..8  | 4..4  | 321003 | 2        | PEMAL4P3  | AL035476    | Plasmoidiu |
| c          | 44    | 64..8  | 4..4  | 1141   | 6        | AX083744  | AX083744    | Sequence   |
| c          | 45    | 64..8  | 4..4  | 5884   | 6        | ATX47067  | AX347067    | Sequence   |

|            |           |
|------------|-----------|
| RESULT 1   | REFERENCE |
| AX154641   | AUTHOR    |
| LOCUS      | TITLE     |
| DEFINITION | TOPIC     |
| ACCESSION  |           |
| VERSION    |           |
| KEYWORDS   |           |
| SOURCE     |           |
| ORGANI     |           |

| ATURES   |  | Sundstrom, Paula (US)<br>Location/Qualifiers<br>1..1470  |  |
|--|--|--|--|
| source   | organism="Candida albicans"<br>/db_xref="taxon:5476"           | IGIN   | 509 a ..260 c 186 g 515 t  |
| SE COUNT   | 1  | Query Match  | 100.0% Score 1470; DB 6; Length 1470;                                      |
| Best Local Similarity  | 100.0% Pred. No. 2e-225;                                       | Mismatches 0; Indels 0; Gaps 0;  |  |
| Matches 1470; Conservative 0;                                  |  |  |  |
| 1 GATATTTCTTTCATTCCCTTAACCGATAAGAAAGTGGAAATAAGCTAT 60          | 1 GATATTTCTTTCATTCCCTTAACCGATAAGAAAGTGGAAATAAGCTAT 60          | Db 901 TACTATAAGCTCAACAAATTACTTCATAAAATGTTATAATTAAACAGTCATCTATAAT 960  | QY 961 TCTTGGAACTCAAAACAGGAATTCGAAAUCTGAGATAATGTCGACCAATTC 1.020           |
| 61 GATAATGTTGATTTGTTGTAATTCATCAACTTAACTTAAAGTACG 120           | 61 GATAATGTTGATTTGTTGTAATTCATCAACTTAACTTAAAGTACG 120           | Db 961 TCTTGGAACTCAAAACGGATTCGAAATCTGAGTTATACTCTAAAGCTAT 1020  | QY 1021 ATTGTAAGAAAGGGAGATTGGTAGGCTCATATACTCTAAAGCTATAAGT 1080             |
| 121 TTGTTGTTGCTCGTCGTCGTTGAGGTTAAACAGGAATAACAG 180             | 121 TTGTTGTTGCTCGTCGTCGTTGAGGTTAAACAGGAATAACAG 180             | Db 1021 ATGTAAGAAAGGGAGATTGGTAGGCTCATATACTCTAAAGCTAT 1080  | QY 1081 TCTTGGAACTCAAAACCTTCTAAACCTATAATAACCTATAATGCTCAACAGCG 1140         |
| 121 GATTTTAAACATAGCAACTTAACTTAACTTAAAGTACG 120                 | 121 GATTTTAAACATAGCAACTTAACTTAACTTAAAGTACG 120                 | Db 1081 TCTTGGAACTCAAAACCTTCTAAACCTATAATAACCTATAATGCTCAACAGCG 1140   | QY 1140 TCTTGGAACTCAAAACCTTCTAAACCTATAATAACCTATAATGCTCAACAGCG 1140         |
| 181 GAAACCTCTCAAAAAAATTGGACCTTACCGCACAATATTGGGATAAACCTGC 240   | 181 GAAACCTCTCAAAAAAATTGGACCTTACCGCACAATATTGGGATAAACCTGC 240   | Db 1140 GATAAGTTAGTTAGCCAGCTGTTTGGCTTATATGACTACATTGTT 1200   | QY 1141 GATAAGTTAGTTAGCCAGCTGTTTGGCTTATATGACTACATTGTT 1200                 |
| 181 GATTTTAAACATAGCAACTTAACTTAACTTAAAGTACG 120                 | 181 GATTTTAAACATAGCAACTTAACTTAACTTAAAGTACG 120                 | Db 1141 GATAAGTTAGTTAGCCAGCTGTTTGGCTTATATGACTACATTGTT 1200   | QY 1201 CACTTTGTTGGACTTAACTGGTTGGAACTCTCCTGATCTGATCTGC 1260                |
| 241 CATAATAAAACTCTTGAACATACGATATGTTATCTTATAACTGGATAATT 300     | 241 CATAATAAAACTCTTGAACATACGATATGTTATCTTATAACTGGATAATT 300     | Db 1201 CACTTTGTTGGACTTAACTGGTTGGAACTCTCCTGATCTGATCTGC 1260  | QY 1260 CTTTTTAAACATAGCAACTTCTGTAAGTCTGTTTCCTCCACTATTTATCATTCT 1320        |
| 241 CATAATAAAACTCTTGAACATACGATATGTTATCTTATAACTGGATAATT 300     | 241 CATAATAAAACTCTTGAACATACGATATGTTATCTTATAACTGGATAATT 300     | Db 1261 CCTTTTAAACATAGCAACTCTGTAAGTCTGTTTCCTCCACTATTTATCATTCT 1320   | QY 1321 TGAAATATGTTAATCAGATAAGTTTCAAAATAACTATAAAATAACGGTCAAATAACCGGCT 1380 |
| 301 GCTTTTTAACATTATGACAAATTGAAAGGAAATGAAAGGTAAAGGTGTCC 360     | 301 GCTTTTTAACATTATGACAAATTGAAAGGAAATGAAAGGTAAAGGTGTCC 360     | Db 1321 TGAAATATGTTAATCAGATAAGTTTCAAAATAACTATAAAATAACGGTCAAATAACCGGCT 1380   | QY 1381 ATTTCAATTCCATTCAAACTGTGTTCTCAAAATAACGGTCAAATAACCGGCT 1440          |
| 301 GCTTTTTAACATTATGACAAATTGAAAGGAAATGAAAGGTAAAGGTGTCC 360     | 301 GCTTTTTAACATTATGACAAATTGAAAGGAAATGAAAGGTAAAGGTGTCC 360     | Db 1381 ATTTCAATTCCATTCAAACTGTGTTCTCAAAATAACGGTCAAATAACCGGCT 1440  | QY 1441 TAGTCAGCCTTGTAGTTTCGTCAATAATG 1470                                 |
| 361 TAACCATGAAATATAAGCTTAAGTTTCTCTGATGGTTACTAAAGGAAATAA 420    | 361 TAACCATGAAATATAAGCTTAAGTTTCTCTGATGGTTACTAAAGGAAATAA 420    | Db 1441 TAGTCAGCCTTGTAGTTTCGTCAATAATG 1470   |  |
| 421 CAAAGTTATAGCGATAACCTGGTAACACAAAATATTGTCAGGTAGCT 480        | 421 CAAAGTTATAGCGATAACCTGGTAACACAAAATATTGTCAGGTAGCT 480        | RESULT 2<br>CAU64206   | Db 901 TACTATAAGCTCAACAAATTACTTCATAAAATGTTAAACAGTCATCTATAAT 960            |
| 421 CTCTCACAGTGAACCTAAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 540 | 421 CTCTCACAGTGAACCTAAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 540 | LOCUS CAU64206   | QY 961 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1020                        |
| 481 CTATGAAATATACAAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 540    | 481 CTATGAAATATACAAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 540    | DEFINITION Candida albicans hyphal wall protein 1 (HWP1) gene, complete cds.   | QY 1021 ATGTAAGAAAGGGAGATTGGTAGGCTCATATACTCTAAAGCTAT 1080                  |
| 481 CTATGAAATATACAAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 540    | 481 CTATGAAATATACAAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 540    | ACCESSION U64206   | QY 1081 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1200                       |
| 541 CTCTCACAGTGAACCTAAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 600 | 541 CTCTCACAGTGAACCTAAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 600 | VERSION U64206..1  | QY 1140 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1200                       |
| 541 CTCTCACAGTGAACCTAAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 600 | 541 CTCTCACAGTGAACCTAAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 600 | TITLE GI:4028879   | QY 1201 CACTTTGTTGGACTTAACTGGTTGGAACTCTCCTGATCTGATCTGC 1260                |
| 601 TAGATTAGTTGTTCTCTTCACTTACCAATTGCAATGCTAGCTTATTG 660        | 601 TAGATTAGTTGTTCTCTTCACTTACCAATTGCAATGCTAGCTTATTG 660        | KEYWORDS SOURCE  | Db 1261 CCTTTTAAACATAGCAACTTCTGTAAGTCTGTTTCCTCCACTATTTATCATTCT 1320        |
| 661 TTCTATTAATGCACTTAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 720  | 661 TTCTATTAATGCACTTAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 720  | ORGANISM Candida albicans  | QY 1321 TGAAATATGTTAATCAGATAAGTTTCAAAATACTATAAAATAACGGTCAAATAACCGGCT 1380  |
| 661 TTCTATTAATGCACTTAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 720  | 661 TTCTATTAATGCACTTAACTTAACTCCTAAGGAATTCTCTATATAAGAAATCC 720  | Eukaryota; Fungi: Ascomycota; Saccharomyces; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.                    | Db 1381 ATTTCAATTCCATTCAAACTGTGTTCTCAAAATAACGGTCAAATAACCGGCT 1440          |
| 721 GTCTCTTGTGTTCTCTTCAACAAATGTTGCAATTCTTAAATAATTGAAATA 780    | 721 GTCTCTTGTGTTCTCTTCAACAAATGTTGCAATTCTTAAATAATTGAAATA 780    | REFERENCE 1 (bases 1 to 2682)  | QY 1441 TAGTCAGCCTTGTAGTTTCGTCAATAATG 1470                                 |
| 721 GTCTCTTGTGTTCTCTTCAACAAATGTTGCAATTCTTAAATAATTGAAATA 780    | 721 GTCTCTTGTGTTCTCTTCAACAAATGTTGCAATTCTTAAATAATTGAAATA 780    | AUTHORS Staab, J.-F., Ferrer, C.A. and Sundstrom, P.   | Db 1441 TAGTCAGCCTTGTAGTTTCGTCAATAATG 1470                                 |
| 781 AACACATAAACTTGTAGTATACTCAACTTATGACTGTTGAAAGTAAGAT 840      | 781 AACACATAAACTTGTAGTATACTCAACTTATGACTGTTGAAAGTAAGAT 840      | TITLE Developmental expression of a tandemly repeated, proline- and glutamine-rich amino acid motif on hyphal surfaces on Candida albicans | QY 1501 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1200                       |
| 781 AACACATAAACTTGTAGTATACTCAACTTATGACTGTTGAAAGTAAGAT 840      | 781 AACACATAAACTTGTAGTATACTCAACTTATGACTGTTGAAAGTAAGAT 840      | JOURNAL J. Biol. Chem. 271 (11), 6298-6305 (1996)  | Db 1501 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1200                       |
| 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | PUBLMED 8626424  | QY 1581 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1200                       |
| 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | REFERENCE 2 (bases 1 to 2682)  | Db 1581 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1200                       |
| 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | AUTHORS Staab, J.-F. and Sundstrom, P.   | QY 1661 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1200                       |
| 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | TITLE Direct Submission  | Db 1661 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1200                       |
| 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | JOURNAL Submitted (18-JUL-1996) Medical Microbiology & Immunology, Ohio State University, 333 West 10th Avenue, Columbus, OH 43210, USA    | Db 1661 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1200                       |
| 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | 841 CAATTTTCTCTTCACTTACATCAACTTACATCAACTTACATCTAC 900          | REFERENCE 4 (bases 1 to 2682)  | Db 1661 TCTTGGAACTCAAAACGGATTCGAACTTCTGAGTTACCT 1200                       |

| TITLE   |        | Submitted (17-DEC-1998) Medical Microbiology & Immunology, Ohio State University, 333 West 10th Avenue, Columbus, OH 43210, USA  |  |
|---|--------|--|--|
| JOURNAL   |        | Sequence update by submitter   |  |
| REMARK  |        | On Dec 17, 1998 this sequence version replaced gi:1915978.   |  |
| FEATURES  |        |  |  |
| source  |        |  |  |
| gene  |        | YQ 1446 ACTCGCTTTAGTTAGTCGCAATATG 1470<br>Db 481 ACTCGCTTTAGTTAGTCGCAATATG 505   |  |
| mRNA  |        |  |  |
| CDS   |        | RESULT 3<br>AF001978 LOCUS AF001978 2188 bp DNA linear PLN 24 -JUL-1997<br>DEFINITION Candida albicans ECE2 gene, complete cds.<br>ACCESSION AF001978<br>VERSION AF001978.1 GI:2275335<br>KEYWORDS Candida albicans.<br>ORGANISM Candida albicans<br>Saccharomyces; mitosporic Saccharomyctales; Candida.<br>SOURCE Sharkey,L.J., Saporito-Irwin,S.M. and Fonzi,W.A.<br>REFERENCE 1. (bases 1 to 2188)<br>AUTHORS Sharkey,L.J., Saporito-Irwin,S.M. and Fonzi,W.A.<br>TITLE BCE2, a differentially expressed gene of Candida albicans<br>JOURNAL Unpublished<br>AUTHORS Sharkey,L.J., Saporito-Irwin,S.M. and Fonzi,W.A.<br>TITLE 2 (bases 1 to 2188)<br>JOURNAL Direct Submission<br>SUBMITTED (01-MAY-1997) Microbiology and Immunology, Georgetown University, 3900 Reservoir Road NW, Washington, DC 20007-2197, USA<br>COMMENT On Jul 24, 1997 this sequence version replaced gi:2209195.<br>FEATURES Location/Qualifiers<br>1. 2188<br>source<br>/organism="Candida albicans"<br>/strain="SC5314"<br>/db_xref="Taxon: 5476"<br>/chromosome="4"<br>/gene="ECE2"<br>/gene_id="AA64014_1"<br>/note="differentially expressed in relation to the extent of cell elongation"<br>/codon_start=1<br>/transl_table=12<br>/protein_id="AA64014_1"<br>/db_xref="GI: 2275336"<br>/translation="MLISTQLIAIAYMISIGATVPOVQDGQGETEBAIQLQKRSYDYY<br>QEPPDYYPOQQQQEQPCDYPQQQQEQPCDYPQQQEQPCDYPQQQEQPCDYPQQQ<br>PCDYPQQQEQPCDYPQQQEQPCDYPQQQEQPCDYPQQQEQPCDYPQQQEQPCDYPQQ<br>PCDYPQQQEQPCDNPQDPODQPCDNPQDPODQPCDNPQDPODQPCDNPQDPO<br>DQDPPDNEPILPNPNTDNPNTDNPNTDNPNTDNPNTDNPNTDNPNTDNPNTD<br>SSTPKPSASTPESSPTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT<br>SSAPATTPNSVTTKSSPLTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT<br>VTTTGTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT<br>EVTTGVVVTTSEEVVTTTFCPTTATTTTATTTTATTTTATTTTATTTTATTT<br>APKSDVTAESAPVTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT<br>TKPAPKSSAPATEPSPVAPGTTTATTTTATTTTATTTTATTTTATTTTATTT<br>SGATTTPESSVTTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCA<br>AATGAAFLI" ATGAAFLI"<br>CDS 759 a 678 c 370 g 875 t<br>ORIGIN 759 a 678 c 370 g 875 t |  |
| BASE COUNT  |        |  |  |
| ORIGIN  |        |  |  |
| Query Match   | 34.48; | Score 505; DB 8; Length 2662;  |  |
| Best Local Similarity   | 100.0% | Pred. No. 1.8e-71;   |  |
| Matches   | 505;   | Conservative 0; Mismatches 0; Indels 0; Gaps 0;  |  |
| QY 966 CGATCCAAAACAAAGGAATTCGGAATTCTGACGATAAATGTCGACTCACAAATCATGTT 1025<br>Db 1 GGATCCAAAACAAAGGAATTCGGAATTCTGACGATAAATGTCGACTCACAAATCATGTT 60    |        |  |  |
| QY 1026 AAAANGGGAGAGTTTGGTAGGGCICATAATCGCTTATAATGTCCTCTAAAGTAATCTAA 1085<br>Db 61 AAAANGGGAGAGTTTGGTAGGGCICATAATCGCTTATAATGTCCTCTAAAGTAATCTAA 120 |        |  |  |
| QY 1086 AACARACACACCTTCTAAACCTATATAAACCTTAATGCTCAACCGGATAA 1145<br>Db 121 AACAAACACACCTTCTAAACCTATATAAACCTTAATGCTCAACCGGATAA 180                  |        |  |  |
| QY 1146 GTTAGTTAGCCAGCTGTTTTTGCCTTATTTTGCCTTATTTTGCCTTACTT 1205<br>Db 181 GTTAGTTAGCCAGCTGTTTTTGCCTTATTTTGCCTTACTT 240                            |        |  |  |
| QY 1206 TTGTGTTGCACTTAAATACGGTTCTCTTGCACACTCTCTCTGCTATCGCCCTT 1265<br>Db 241 TTGTGTTGCACTTAAATACGGTTCTCTCTGCTATCGCCCTT 300                        |        |  |  |
| QY 1266 TTAAACATAGCAACTCTGTAAGCTTCTCTTGCACCTCTCTGCTATCGCCCTT 1325<br>Db 301 TTAAACATAGCAACTCTGTAAGCTCCTTCTCTCTGCTATCGCCCTT 360                    |        |  |  |
| QY 1326 TATGTAATCAGATAATGTTTCAAAAACATAATAAACGGTCAAATACCGCTTATT 1385<br>Db 361 TATGTAATCAGATAATGTTTCAAAAACATAATAAACGGTCAAATACCGCTTATT 420          |        |  |  |
| QY 1386 CAATTCCACACTCAACTCTGTTCTCAACATATAACCAAGGAATCTCTATAGTC 1445<br>Db 421 CAATTCCACACTCAACTCTGTTCTCAACATATAACCAAGGAATCTCTATAGTC 480            |        |  |  |
| QY 1448 CTATAGTCACCTCGCTTTAGTTCTGTCATATG 1470<br>Db 121 CTATAGTCACCTCGCTTTAGTTCTGTCATATG 153  |        |  |  |

RESULT 4  
 CAC35A5/c  
 LOCUS C. albicans cosmid Ca35A5.  
 DEFINITION C. albicans cosmid Ca35A5.  
 ACCESSION AL033396  
 VERSION AL033396.1  
 KEYWORDS ARS; Can1; CDP-alcohol phosphatidyltransferases; chs1; cytochrome P450; DNA polymerase delta; histidine kinase; isocitrate dehydrogenase pseudogene; LTR; pol3; Rcc1; regulator of chromosome condensation; rehydrin; RPS1-like region.  
 SOURCE Candida albicans  
 ORGANISM Candida albicans  
 Eukaryota; Fungi; Ascomycota; Saccharomyctina; Saccharomycetes; Saccharomyctales; mitosporic Saccharomycetales; Candida.  
 1. (bases 1 to 42565) TAIT, E., SIMON, M.C., KING, S., BROWN, A.J., GOW, N.A. and SHAW, D.J.  
 A *Candida albicans* genome project: cosmid contigs, physical mapping, and gene isolation  
 JOURNAL *Fungal Genet. Biol.* 21 (3), 308-314 (1997)  
 MEDLINE 97290243  
 PUBLMED  
 REMARK Article No. FG970983  
 2 (bases 1 to 42565) Oliver, K. and Harris, D.  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 42565) Direct Submission  
 AUTHORS barrellesanger.ac.uk  
 Submitted (05-NOV-1998) On behalf of the pilot sequencing project on the *Candida albicans* strain 1161 genome. Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: barrellesanger.ac.uk  
 Cosmids supplied by Prof. Duncan Shaw, [3] Department of Molecular and Cell Biology, The Institute of Medical Science, University of Aberdeen, Foresterhill, Aberdeen, AB25 2ZD, UK  
 Notes: Funding: sequencing funded by Beowulf Genomics Ltd. CDS are numbered using the following system eg CAC20C1.01c. CA (C. albicans), 20C1 (cosmid name).  
 01 (first CDS), c (complementary strand).  
 The more significant matches with motifs in the PROSITE/PPFAM database are also included but some of these may be fortuitous.  
 IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid Ca35A5 is likely to map to region R, chromosome 7.  
 Location/Qualifiers  
 1. 4.2365 /organism="Candida albicans"  
 /strain="1161"  
 /db\_xref="Taxon:5476"  
 /map="Chromosome 7"  
 /clone="Cosmid Ca35A5"  
 /complement(1..785)  
 /gene="Ca35A5\_01c"  
 /complement(-1..785)  
 /note="Ca35A5\_01c", partial orf, len: > 261 aa, most similar to *RPT1*, YEAST ethanaminephosphotransferase (EC 2.7.8.1) (301 aa), fasta scores opt: 984, E1: 0, (56.3% identity in 263 aa overlap), contains PS00379 CDP-alcohol phosphatidyltransferases signature"  
 /codon\_start=1  
 /transl\_table=12  
 /product="putative alcohol phosphatidyl transferase"  
 /protein\_id="Ca35A5\_01c"  
 /db\_xref="GI:3850144"  
 /note="MGI:IPINLQNLKLYKSSSEHSHSISLKKWWNNEFVQFPL  
 /SNAPNVVTLIGFFILGNNLMTYFYDQPLNTEQFGLMYCQDFGCGCH  
 /ARTGQSGPLGFFLDFDISDALNTLGLTFVAVSLRMGGLLISOFASVCFNFTSTW  
 /EXYHTHTLFLSSFSQGPVEGLMICIVYITRIGFGDIDWLFELNLTSGYGYKV  
 misc\_feature  
 complement(423..491)  
 /gene="Ca35A5\_01c"  
 /signature="4307..11124  
 /note="RPS1-like region"  
 complement(7606..8085)  
 /note="region showing similarity with isocitrate dehydrogenase"  
 13175..14713  
 /gene="Ca35A5\_02"  
 13175..14713  
 /gene="Ca35A5\_02"  
 /note="Ca35A5\_02, unknown, len: 512 aa, possible glycoprotein"  
 /codon\_start=1  
 /transl\_table=12  
 /product="hypothetical protein"  
 /protein\_id="CAA21945.1"  
 /db\_xref="GI:3850145"  
 /translation="MSEVPNOEPSSSTLKKQNSASSTTKVIGAPSSQQLANDENTP  
 KLNSSNTSSSSNTLNNTSSNTSNSNLSPLTRKNSLNSNTSVPSSDDSSPKTLESP  
 SESSRSVSNRLRAKKLFSKNQNSRKPDLFDTSV1 SEUDSSRPTKVLGNSLNM  
 TKPKSSNSNSPMSPTTEHSLFHRSHHNSNTLQDRTNOVATPLSGRFDKPSKTT  
 LCQLENSNSNSLISNPELAQTLYNTNPNTSLERETNLDNTNSFLOTHKMLVPAKSF  
 IQNKLINKYHQTEVGLGIYSELDHNDKTYISNLHYLKP1LEPTPSRS1SDSQKSXMR  
 PLSASVSEETANFVKESCLHQINHERRFSRVSYSSLGKDVDFDYLQSLFTRSH  
 KLMSSLHSNLCQTTEPSEYSLQALITNAWKKYNAVREYLLSFOQYIQLNPFQY  
 NGSKLIRIDLLVLSFRKVFTIEQGIGSGERETSQFLGNAESEDLTGNGLLTSLAVL  
 SSIS"  
 gene  
 18031..18456  
 /gene="Ca35A5\_03"  
 CDS  
 18031..18456  
 /gene="Ca35A5\_03"  
 /note="Ca35A5\_03, unknown, questionable orf, len: 141 aa"  
 /codon\_start=1  
 /transl\_table=12  
 /product="questionable orf"  
 /protein\_id="CAA21946.1"  
 /db\_xref="GI:3850146"  
 /translation="MNQSNMNSKNTERRNSNTEDDLYHYHEYVSLSVQHEQNQKQELS  
 MVEIERLKKEANRSDIKDIIKILIVNDQKIDSQNAFSSLMKSALKQDKRNQYRKL  
 19043..19210  
 /note="BLASTN hit (score: 669, strand: +1), range: 8-175 to AF069450 C. albicans retrotransposon long terminal repeat zeta"  
 /note="BLASTN hit (score: 669, strand: +1), range: 168-24 to AF069450 C. albicans retrotransposon long terminal repeat zeta"  
 misc\_feature  
 20359..20769  
 /note="BLASTN match (score: 1704, strand: +1), range: 49-459 to CAAURESE X65035 C. albicans DNA of an autonomously replicating sequence (ARS)"  
 20719..21001  
 /note="BLASTN match (score: 1298, strand: +1), range: 405-687 to CAAURESE X65035 C. albicans DNA of an autonomously replicating sequence (ARS)"  
 20808..23888  
 /note="Ca35A5\_04, Chs1 gene, len: 1026 aa, CHS1\_CANAL  
 /gene="Chs1"  
 /gene="Ca35A5\_04"  
 /gene="Chs1"  
 /gene="Ca35A5\_04, Chs1 gene, len: 1026 aa, CHS1\_CANAL  
 /note="99.9% identity in 776 aa overlap) but open reading frame here extends 250 aa upstream of CHS1\_CANAL start at 21561"  
 /codon\_start=1  
 /transl\_table=12  
 /product="chitin synthase"  
 /protein\_id="CAA21947.1"











|    |       |   |       |    |       |  |       |
|----|-------|---|-------|----|-------|--|-------|
| Qy | 224   | ATTCGGATAAACTTGGCCATAATAAACCTTGTAAACATACGATATGTTATCTTTTC  | 283   | Db | 47616 | CATGATTAATAATTCTATTATTATTTTATTATTTTATTATTTTCAACCCAC    | 47557 |
| Db | 48696 | ATTCGAAANAAAAMAAAAGAGAAATTAACATACGATATGTTATGTTCAACT       | 48537 | Qy | 1353  | CTATIAAATACGGICAAATAACCGGCPATTTCAACTTCATCAACTGTTTCCTAA | 1412  |
| Qy | 284   | ATTAATGGAAATATTTGGCTTAACTTGTAACTTGTAACTTGTAACTTGTAACT     | 343   | Db | 47556 | ATGTTAAATAATAATAGCAACACGGCAATTTGTTAAATAATTTTGTATTTTT   | 47497 |
| Db | 48636 | ATTCCTTTGGGGTGAATGGGAAATGGGAAATGGGAAATGGGAAATGGGAAATGG    | 48577 | Qy | 1413  | CAATAT 141.8   |       |
| Qy | 344   | AAAGCTTAAAGGTGCTAACCTTGTAAATTAAGCTTGTAACTTGTGGTT          | 403   | Db | 47496 | CAATT 47491  |       |
| Db | 48576 | ATTAATGGTAAATTTGGAAATGGAAATGGAAATGGAAATGGAAATGGAAATGG     | 48517 |    |       |  |       |
| Qy | 404   | ACTTAAAGGAAATAACAAACTTAACTGGATAACCTGGCTTAAGGTGTAAACAAATA  | 463   |    |       |  |       |
| Db | 48516 | ATTAATGGTTAAATTTGTAAATTAATTAATTAATTAATTAATTAATTAATTA      | 48457 |    |       |  |       |
| Qy | 464   | TATTGGCACGTAGCTCTATGAAATAACAAACTTAACTGGCTTCCTCTA          | 523   |    |       |  |       |
| Db | 48456 | TGTTAAATTTGTAGTGTGGTGAATGGAATCAAGATAATAATGGGGCATTATTATA   | 48397 |    |       |  |       |
| Qy | 524   | TATTAATAGGAATCCCTCTCACAGTGAACGAACTTATCCATGAAATTATCAGTCCAC | 583   |    |       |  |       |
| Db | 48396 | TTTTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAACAAAA | 48337 |    |       |  |       |
| Qy | 584   | TAATTCACAAATAAAATAGATAGTGTATTGTTCTTCAGTACAATTACTACATTAT   | 643   |    |       |  |       |
| Db | 48336 | ATAAAACTAAAGAAATAAAATAAAGTGTATTGTTTACATGGTTATGGTTAAAG     | 48277 |    |       |  |       |
| Qy | 644   | GCAATGCTAGCTTAATGGTCATAATTAGGCATGTGTGACACCCCTAACTGAACTTA  | 703   |    |       |  |       |
| Db | 48276 | TGAAATAAGGTGTAAATAATAAAATAAAATAAAATAAAATAAAATAAAATAAA     | 48217 |    |       |  |       |
| Qy | 704   | GTATCCCATATTTCCTTGTCTTCTTCTTACAAATAATGTTCCAGAAATTTT       | 763   |    |       |  |       |
| Db | 48216 | TTGGATAA        | 48157 |    |       |  |       |
| Qy | 764   | TAAAAATATTGAAAAAACACATAACACTTGTGATGATAATACTTGTACTTG       | 823   |    |       |  |       |
| Db | 48156 | AAATATAATAATAAAACATGAAAGAAAATAAAATAAAATAAAATAAAATAAA      | 48097 |    |       |  |       |
| Qy | 824   | TTTGTAAAGTAAAGATCAAATTTCTAACTCGACTTAATGCACTTACATCAACTGG   | 883   |    |       |  |       |
| Db | 48096 | GAATGTGTGAAGCAGAACCTTAACTTTAAATTTAAATTCGGTTATTATTTTA      | 48037 |    |       |  |       |
| Qy | 884   | ATGTTATTGGCATCTACTAACTTAACTTAACTTAACTTAACTTAAATGTTAT      | 939   |    |       |  |       |
| Db | 48036 | TTTATTTATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT      | 47977 |    |       |  |       |
| Qy | 940   | ATAATCAAGCTCATCTATAATTCTTGGATCCAAAACACGAAATTCTGTGACG      | 999   |    |       |  |       |
| Db | 47976 | TGATAATTTATATAATAATAATAATAATAATAATAATAATAATAATAATA        | 47917 |    |       |  |       |
| Qy | 1000  | ATAATGTGGCTACAATTCTGTTAAAGGGAGGTGTTGAGCTTAACTACGTTCT      | 1059  |    |       |  |       |
| Db | 47916 | AAAAAAAGAAAAAAAGAAAAAAAGAAAAAAATCAAAAAAAAGAAAAAAAGAAAA    | 47857 |    |       |  |       |
| Qy | 1060  | TATAATGTACCTCTAAAGTAATCTAAACACACACACCTTCTAAACCTATAATA     | 1119  |    |       |  |       |
| Db | 47856 | AAATCCTTTGGCAGGAAATTGTGTTAAATTTGAAACTTGTATTTGAAATTTA      | 47797 |    |       |  |       |
| Qy | 1120  | ACCC-----TAATGGCTACAACCGGATAATGGTGTGCTTGTGTTGGGAAATAATTG  | 1172  |    |       |  |       |
| Db | 47796 | ACCGTTTTTGATGACTACATTGGCTACTTGTGTTGGCTACTTGTGTTGGT        | 1232  |    |       |  |       |
| Qy | 1173  | GCCTTATTGTTATGACTACATTGGCTACTTGTGTTGGCTACTTGTGTTGGT       | 1232  |    |       |  |       |
| Db | 47736 | TTTTTTTTGATTTTGTCCATTGTTCCATTGTTCCATTGTTCCATTGTTCCATT     | 47677 |    |       |  |       |
| Qy | 1233  | CAACTCTCTTGTATCACCTGTTAACATGCAACTCTGGCTTAACTGAAAGTGC      | 1292  |    |       |  |       |
| Db | 47676 | AGATT     | 47617 |    |       |  |       |
| Qy | 1293  | CTTCCTTTCCACTTATCATCTGGAAATACTGAAATGTAATCAGATACTGTT       | 1352  |    |       |  |       |





|   |  |                                 |
|---|--|---------------------------------|
| /codon_start=1  | Db 145856 ATAGCATATTAATAGATAATAAACTATAATTATTTGGTAAAAAATATT 145915  |                                 |
| /product="hypothetical protein, PFC0715c"                                     | Qy 988 GAAATCTGAGATGAAATGTCGACTCACATTCACTGAAAAGGAGACTTTGGTAGC 1047   |                                 |
| /protein_id="CAB1147_1"   | Db 145916 AAAAAAATATTAAATTAGAAATTAATTAAAGAAATTAATTGATAAAAGTTG 145975   |                                 |
| /db_xref="SPTRMEL:077391"   | Qy 1048 CTCATATCGCTTAATATGCCCTAAAGTAATCTAAACACACCTTCTAAAA 1107   |                                 |
| /translation="MANTLYIFIGIIVILCCSRCCNNVNNENIGKSHYPTFLKERENRN                   | Db 145976 ACATAAAGGGGAACTAAACAAACACACCTTCTAAAA 1107  |                                 |
| NNFCNSVNNIGRKKIKKKNNVPHSGYIILPLANKLSSNTNNNSRKICLVDYYP                         | Qy 1108 CCTCATATTAACCCCTAAAGGCTGACAACCCGGATAAGTTAGGCCAGCTGTTTT 1167  |                                 |
| NEKKOGNNTLEHHVHTNKMKSKAERKNNKVKRNVMDKNGDI                                     | Db 146036 CTCATATTCGCTGGTTAAACAAAGTACATTAAATTAATATATGATA 146035  |                                 |
| LEPAMKTAEEYTKGKEDILIBREKEKKKVEKAKDDEWDFTDKLDRYIDMKYV                          | Qy 1168 TTTCCTCCCTTA-<br>-TTTTTATGACTCATTTGTCACATTGTCACCTTTGTCACCTTAATAC 1225  |                                 |
| YRNNTIYRTSHLNVNCWMMRNSVNDKINNIEQREKKEKDEEJKRKKEKE                             | Db 146096 ATTTCCTAAAGCTTTTGTGTTTTCTCATATTCCCTTGTGTTAGCAGAC 146155  |                                 |
| L-NLISDNEKDTHNETNINYKVNVNDKYNHDELIKGRKAKETDDEQSELKEKD                         | Qy 1226 GTTTTGCAACTCTCTCTCTATCACCTGTATGCCCTTTAACATAGCAACTCTGT 1285   |                                 |
| TSINYSKDRNQYLNPESEDDNEKUNSEMSLKLXNNKNSKYESLXKGNILKOLV                         | Db 146156 ATTTCCTCAATTTCCTCTCTCTCTGATCTTTGCTTGTGCAATT 146215   |                                 |
| TEGEGDLOPKDKEYTERKKKDDINNKDNNYDNKNDYNNLLENNEDKDR                              | Qy 1286 AAAGTCCCTTCTCTCTCCACATTTATCATTCATCTTCAACCATATCTAAATT 1345  |                                 |
| WATLDSNLFLYDEEYLDKERYMKSVNTIDYDSVNLNNKNTSTMELFPL                              | Db 146216 TCTCTCAATTCTGTTCTGTTGATATCATCTTCAACCATATCTAAATT 146275   |                                 |
| IDSHTLIOKCLTUMKLEKAYIKNKNLFLFSTSDNPINKDNEYDYLINENEWIK                         | Qy 1346 TCAAAACTATAAT 1360   |                                 |
| MELYLCNUYHNEKMGKLTPMNNEHKTNNELKNMDSQNNMILGDSNKGKV                             | Db 146276 ACATATCTCTAACT 146290  |                                 |
| DNTFKENIINKLUSDKDALRKNSAEESESSKKNVTKNVLNNNSQVNLNDEQI                          |  |                                 |
| VENVENTKVNENVKVNENVKVNENVKVNENVKVNENVKVNENVKVNENVKVN                          |  |                                 |
| KELETKTSEDFKLKLENNVKVNENVKVNENVKVNENVKVNENVKVNENVKVN                          |  |                                 |
| SKCVDNIKKRKSQMKGDIGKLEGRTYIYENAFSFSLYFFRPLKKHLDI                              |  |                                 |
| DNNEFFSLTINKHNTIRDNFNPINSDSITSLTDNEDNHMENEENEMQFFPYDITNI                      |  |                                 |
| Query Match 5.0%  | Best Local Similarity 4.9%   | Score 73.4; DB 3; Length 16439; |
| Matches 470; Conservative 0; Mismatches 621; Indels 4; Gaps 2;                | Pred. No. 0, 0.0051;   |                                 |
| Qy 268 ATAGTTATCTTCTTCTAACTGGAATATTGCTTTTAACATTGAAACATATG 327                 |  |                                 |
| Db 145198 ATATATATATATATATATATATATCTCTTTTTCTTAACAAATTAAAGGGAC 145257          |  |                                 |
| Qy 328 AAAAAGGAAATGAAAGGTAAGACTGCTAACATTGAAATAATAGCTTAAGGT 387                | RESULT 11  |                                 |
| Db 145258 ATGCAAAAGAAATGCAAGTCAGTAATGAGATACTGAGATCT 145317                    | PFMA13P4   |                                 |
| Qy 388 TTTCCTCTGATGGGTTAACTAAAGGAATAACAAAGTTAACCTGCGTA 447                    | LOCUS  |                                 |
| Db 145318 ATATCCAAAGTCTTAATATGAAGTAATACATACATATGCTTTTATACACGCA 145377         | PFMA13P4   |                                 |
| Qy 448 AGGTGTCAACAAATAATTGCAATTGCAATTGAGCTTACAGTAAATCCAT 507                  | DEFINITION   |                                 |
| Db 145378 CATATATATATATATATATATATATATATATATATGAGATCT 145437                   | PFMA13P4   |                                 |
| Qy 508 TAGGAATTCCCTCTATATATAGGAATTCACAGTGAACTGAAATTCCATC 567                  | ACCESSION  |                                 |
| Db 145438 CACATATGATATCATCTACATATGAAAGTAATACATACATATGCTTTTATACACGCA           | AL008970_3   |                                 |
| Qy 568 TGAATTACAGCCACAACTCATCAATAAAATGATTAGTGTGATCTCTCAGTA 627                | VERSION  |                                 |
| Db 145498 GAAGACATTATGATATGATATGATATGATCTTCCCTGTGAATA 145555                  | HTG.   |                                 |
| Qy 628 CAAATTACATTATGAACTGCTTAATGCTTACATTGCTTACATTGCTTAC 687                  | KEYWORDS   |                                 |
| Db 145556 AAGTTTTTTTTTATTATTTTAAAGTAATTTAATTTACATARA 145615                   | Plasmidium falciparum 3D7.   |                                 |
| Qy 688 AATTGCAACATTAACGTGATCCATTTGCTTAATGCTTACATTGCTTAC 747                   | Plasmidium falciparum 3D7.   |                                 |
| Db 145616 ATTATATATATATAGGTATTACATTGCTTACATTGCTTAC 747                        | Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.   |                                 |
| Qy 748 ATGTTCCAGAATTTTAAAGGATTTGAAAAACACATAACACTTGTGATGATA 807                | 1 (bases 1 to 113880)  |                                 |
| Db 145676 GTGATCTTTTATTATTTATTTATTTATTTATTTATTTATTTAT 145735                  | REFERENCE  |                                 |
| Qy 808 ATCAACTATGACTGTGTTAAAGTAAATGTTCTGCTCTTCTGCTTAC 867                     | AUTHORS  |                                 |
| Db 145736 ATATATATATCACTGAAAGAAAAACATATATATATATATATA 145795                   | Bowman, S., Lawson, D., Basham, D., Chillingworth, T.,   |                                 |
| Qy 868 ACTTGTACATCACTGCTGATGTTATGCTATCTACTAATGCTCAAATTCCTT 927                | Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S.,  |                                 |
| Db 145796 TAT 145855                  | Hornby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J.,  |                                 |
| Qy 928 CAAAAATGTTATAATTAACAAAGTCATCTATAATTCTTGGATCCAAAAGGAATTCG 987           | Moule, S., Mungall, K., Murphy, L.-A., Rutter, S., Skelton, J., Squires, R., Squares, S., Sulston, J. E., Whitehead, S., Woodward, J. R., Newbold, C. and                        |                                 |
| Db 145856 ATAGCATATTAATAGATAATAAACTATAATTATTTGGTAAAAAATATT 145915             | Barrell, B.G.  |                                 |
| Qy 988 GAAATCTGAGATGAAATGTCGACTCACATTCACTGAAAAGGAGACTTTGGTAGC 1047            | The complete nucleotide sequence of chromosome 3 of Plasmodium   |                                 |
| Db 145916 AAAAAAATATTAAATTAGAAATTAATTGATAAAAGTTGATAAA 1107                    | falciparum   |                                 |
| Qy 1048 CTCATATCGCTTAATATGCCCTAAAGTAATCTAAACACCTTCTAAAA 1107                  | JOURNAL  |                                 |
| Db 146036 CTCATATTCGCTGGTTAAACAAAGTACATTAAATTTATGATA 146095                   | MEMLINE  |                                 |
| Qy 1108 CCTCATATTAACCCCTAAAGGCTGACAACCCGGATAAGTTAGGCCAGCTGTTTT 1167           | 99376085   |                                 |
| Db 146096 ATTTCCTAAAGCTTTTGTGTTTTCTCATATTCCCTTGTGTTAGCAGAC 146155             | PUBLMED  |                                 |
| Qy 1168 TTTCCTCCCTTA-<br>-TTTTTATGACTCATTTGTCACATTGTCACCTTTGTCACCTTAATAC 1225 | 10448855   |                                 |
| Db 146216 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146275              | REFERENCE  |                                 |
| Qy 1226 GTTTTGCAACTCTCTCTCTATCACCTGTATGCCCTTTAACATAGCAACTCTGT 1285            | 2 (bases 1 to 113880)  |                                 |
| Db 146276 ACATATCTCTAACT 146290   | AUTHORS  |                                 |
| Qy 1286 AAAGTCCCTTCTCTCTCCACATTTATCATCTTCAACCATATCTAAATT 1345                 | Hamlin, N., Bowman, S., Churcher, C., Lawson, D., Quail, M. and  |                                 |
| Db 146277 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146275              | Barrell, B.  |                                 |
| Qy 1346 TCAAAACTATAAT 1360  | UNPUBLISHED  |                                 |
| Db 146278 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | 3 (bases 1 to 113880)  |                                 |
| Qy 1360 TCAAAACTATAAT 1360  | REFERENCE  |                                 |
| Db 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | Lawson, D., Bowman, S. and Barrell, B.   |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | TITLE  |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | JOURNAL  |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | COMMENT  |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | Submitted (18-Nov-1998) P.falciparum Genome Sequencing Consortium, CB10 13a, UK  |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | On Apr 30, 2000 this sequence replaced the gi:4493924.   |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | For more information about this sequence or the Malaria Project, see <a href="http://www.sanger.ac.uk/projects/P_falciparum/">http://www.sanger.ac.uk/projects/P_falciparum/</a> |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | Location/Qualifiers  |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | 1..113880  |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | /organism="Plasmidium falciparum"  |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | /strain="3D7"  |                                 |
| Qy 146290 TCTCTCAATTCTGTTCTTGTGATATCATCTTCAACCATATCTAAATT 146290              | /db_xref="Taxon:36329"   |                                 |





|  |   |   |                                      |                |                 |  |       |
|--|---|---|--------------------------------------|----------------|-----------------|--|-------|
| Db   | 190   | ATTATATATATTATTTATTAAATTATTTAAATAATTATTTAATTATTCATTATAC | 131                                  | Qy             | 472             | ACGTTAGCTCATAGAAAATATAACAAACTAAATCCTTAAGGGATTTCCTCTATATAAT | 531   |
| Qy   | 1012  | TGCAAACTTCACTGTTAAAAGGAGAGTTGGCTCATAAACGCTTAAATGTAACCT  | 1071                                 | Db             | 43490           | TAATACTCATATAATTATTTATATAATTATATAATTATATAATTATATAATTAT     | 43431 |
| Db   | 130   | ATAAATATATATATAATTATTTATTTATTTATTTATTTATTTATTTATTTATTTA | 71                                   | Qy             | 532             | AGGAAATCCCTCTCACAGTGACTGAAATTCCATCTGAAATTCTGACTTAATCCCA    | 591   |
| Qy   | 1072  | CTAAAGTAATCTAAA 1087                                    |                                      | Db             | 43430           | ATTTATAGATTATATATAGATAATAATTATATAATTATATAATTATTTATTTA      | 43371 |
| Db   | 70  | TATATATAATCACACA 55                                     |                                      | Qy             | 592             | TCATAAAATAGATTAGTAGTTAGTGTCTCTAGTACAAATTACATTATGCTATGCT    | 651   |
| RESULT 13  |   |   |                                      | Db             | 43370           | CCATAATAAGCATTAT-CATTTTAAAAAAATGAATAGTAAATAATTATATAT       | 43312 |
| AC005505/C   | AC005505  | AC005505  | 192929 bp DNA linear HTG 12-AUG-2000 | Qy             | 652             | AGCTTATGTCATAATTAGCCATGTTGTCACCCCTTAATTCGAACATACTGTATCCA   | 710   |
| DEFINITION   | Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS   | ***, 1 ordered piece.                                   |                                      | Db             | 43311           | AATTATACATTATATACTACATAGTGTAAAGTAAACGTTATTTTATATATATA      | 43252 |
| ACCESSION  | AC005505  | AC005505  |                                      | Qy             | 711             | TATTTCTCTGCTCTTCTTCTTCTAACTAAATGTTCCAGAAATTCTTAAAAAA       | 770   |
| VERSION  | AC005505.8  | GR-19797710   |                                      | Db             | 43251           | TTAATAATATAATTATTCATCTCTGATCATACTTTTTTTTTAAATTAAT          | 43192 |
| KEYWORDS   | HTGS_PHASE2   |   |                                      | Qy             | 771             | TATTTGAAAAACACATAACACTTGGATTTGATATACATTTGCTTGTGAA          | 830   |
| ORGANISM   | Plasmodium falciparum   |   |                                      | Db             | 43191           | TAACAAACATTAATAAAATAATTTAAATTTAAATTTAAATTTAAATACAA         | 43132 |
| Bukaryota; Alveolae; Apicomplexa; Haemosporida; Plasmodium.        |   |   |                                      | Qy             | 831             | AGTAAAGAAATCAAATTCTTCAACTCTGACTATGCACTACATCAACTGATGTAT     | 890   |
| HYMAN, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T., |   |   |                                      | Db             | 43131           | AAAAATTAACCTATTATATAATTATATAATTATATAATTATATAATTATA         | 43072 |
| Kurd, O.B., Conway, A.B. and Davis, R.W.                           |   |   |                                      | Qy             | 891             | TTGGCATCTACTATAAGCTCAAACAAATTCTTCAAAATGTTATAACAACT         | 950   |
| TITLE  | Plasmodium falciparum   |   |                                      | Db             | 43071           | TTTTTTTTTTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT       | 43012 |
| JOURNAL  | Unpublished   |   |                                      | Qy             | 951             | CATCTATACTCTGTTGATCTGAAACAACTGAAATCTGAGATAATCTGCGA         | 1010  |
| REFERENCE  | 2 (bases 1 to 192929)   |   |                                      | Db             | 43011           | CTTTAAAAAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATGACA       | 42952 |
| AUTHORS  | Hyman, R.W., Qin, F., Fung, E.L., Rowley, D., Mao, J., Tamaki, T.,  |   |                                      | Qy             | 1011            | CTCAGAAATTCATGTTAAAGGGAGAGTTGGTAGGTCTATATCGTTATGTGACC      | 1070  |
| JOURNAL  | Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA |   |                                      | Db             | 42951           | ACATTTTCTCTGAACTATTATTTATTTATTTATTTATTTATTTATTTAT          | 42899 |
| COMMENT  |   |   |                                      | Qy             | 1071            | TCTAAAGTATCTAAACAAACACAAACCTTCTTAAACCTTAAATACCCCTAAATGGC   | 1130  |
| AUTHORS  | Hyman, R.W., Qin, F., Fung, E.L., Rowley, D., Mao, J., Tamaki, T.,  |   |                                      | Db             | 42898           | TAGATAAAAGAAAAAAAGAAAAAAATATATATATATATATATATATATAT         | 42839 |
| JOURNAL  | Direct Submission   |   |                                      | Qy             | 1131            | TCACACCGGATGATGTTAGTTGGCCATTGTTGGCTTATTCTTATGACTA          | 1190  |
| COMMENT  |   |   |                                      | Db             | 42838           | ATAATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT          | 42779 |
| ** NOTE: this is a 'working draft' sequence. It currently          |   |   |                                      | Qy             | 1191            | CATTGTTGTCACCTTTGTCGACTTTAACCGTTTGTGACTCTCTTGTATCA         | 1250  |
| * consists of 1 contigs. Gaps between the contigs                  |   |   |                                      | Db             | 42778           | TTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT        | 42719 |
| * are represented as runs of N. The order of the pieces            |   |   |                                      | Qy             | 1251            | CCTGATCCGCCTTTCATACATAGCAA 1278                            |       |
| * are believed to be correct as given, however the sizes           |   |   |                                      | Db             | 42718           | ATTTTTAATTCTACATTTTGTAAAAA 42691                           |       |
| * of the gaps between them are based on estimates that have        |   |   |                                      |                |                 |  |       |
| * provided by the submitter.                                       |   |   |                                      |                |                 |  |       |
| * This sequence will be replaced.                                  |   |   |                                      |                |                 |  |       |
| * by the finished sequence as soon as it is available and          |   |   |                                      |                |                 |  |       |
| * the accession number will be preserved.                          |   |   |                                      |                |                 |  |       |
| * the accession number will be replaced.                           |   |   |                                      |                |                 |  |       |
| FEATURES   | Location/Qualifiers   |   |                                      |                |                 |  |       |
| source   | 1. 'Plasmodium falciparum'  |   |                                      |                |                 |  |       |
|  | /organism="Plasmodium falciparum"   |   |                                      |                |                 |  |       |
|  | /Db_xref="txon:5833"  |   |                                      |                |                 |  |       |
|  | /chromosome="12"  |   |                                      |                |                 |  |       |
|  | /clone="PFYWC181"   |   |                                      |                |                 |  |       |
|  | /clone="3D7"  |   |                                      |                |                 |  |       |
| BASE COUNT   | 81142 a 17397 c 22111 g 72279 t   |   |                                      |                |                 |  |       |
| ORIGIN   |   |   |                                      |                |                 |  |       |
| Query Match  | 4.8%  | Score 70.4;   | DB 2;                                | Length 192929; |                 |  |       |
| Best Local Similarity  | 43.7%   | Pred. No. 0.015;  | 0; Mismatches 581;                   | Indels 9;      | Gaps 3;         |  |       |
| Matches  | 458; Conservative   |   |                                      |                |                 |  |       |
| RESULTS  | 14  |   |                                      |                |                 |  |       |
|  | AC115598  |   |                                      |                |                 |  |       |
| LOCUS  | AC115598  | 56099 bp  | DNA                                  | linear         | HTG 21-MAR-2002 |  |       |
| DEFINITION   | Dictyostelium discoideum chromosome 2 map 710700-766797 strain AX4,   |   |                                      |                |                 |  |       |
| VERSION  | AC115598.1  | GI:19570017   |                                      |                |                 |  |       |
| KEYWORDS   | SEQUENCING IN PROGRESS **, in ordered pieces.   |   |                                      |                |                 |  |       |
| SOURCE   | HTGS_PHASE2.  |   |                                      |                |                 |  |       |
| ORGANISM   | Dictyostelium discoideum.   |   |                                      |                |                 |  |       |
| REFERENCE  | Eukaryota; Mycetida; Dictyosteliida; Dictyostelium.   |   |                                      |                |                 |  |       |
| AUTHORS  | 1 (bases 1 to 56099);   |   |                                      |                |                 |  |       |
|  | Gloeckner, G., Eichinger, L., Szafranski, K., Pachebat, J., Dear, P.,   |   |                                      |                |                 |  |       |
|  | Lehmann, R., Baumgart, C., Parra, G., Apil, J.F., Guigo, R., Kumpf, K.,   |   |                                      |                |                 |  |       |
|  | Tunggal, B., Cox, E., Quail, M.A., Platzer, M., Rosenthal, A., and  |   |                                      |                |                 |  |       |



| RESULT 1.5 |   | RESULT 1.5 |   |
|------------|---|------------|---|
| VERSION    | KEYWORDS  | VERSION    | KEYWORDS  |
| PFM13P7    | HTG; 40S Ribosomal protein S3A; acyl transferase; | PFM13P7    | HTG; 40S Ribosomal protein S3A; acyl transferase; |
| LOCUS      | 253305 bp DNA linear INV 22-MAY-2000              | LOCUS      | 253305 bp DNA linear INV 22-MAY-2000              |
| DEFINITION | Plasmidium falciparum MAL3P7, complete sequence.  | DEFINITION | Plasmidium falciparum MAL3P7, complete sequence.  |
| ACCESSION  | AL015916  | ACCESSION  | AL015916  |
|            | GACCAAGATTCACACATTTAATTCTTTCAGTTTAAATTATTTT       |            | GACCAAGATTCACACATTTAATTCTTTCAGTTTAAATTATTTT       |
|            | 159716  |            | 159716  |
|            |   |            |   |
| OY         | 653   | OY         | 653   |
|            | GCTTATTGTCATATTAGCCATGTTGCACACCTTAATTGACATTA      |            | GCTTATTGTCATATTAGCCATGTTGCACACCTTAATTGACATTA      |
|            | 712   |            | 712   |
|            |   |            |   |
| Db         | 15976   | Db         | 15976   |
|            | ATTTGGGGTCTGGAAATTAAATTAAATCCTTTCAGGTTAAATTT      |            | ATTTGGGGTCTGGAAATTAAATTAAATCCTTTCAGGTTAAATTT      |
|            | 16035   |            | 16035   |
|            |   |            |   |
| QY         | 713   | QY         | 713   |
|            | TTCATCTTCCTTTCACACATTTGAGTTGATAAATCAACTATTG       |            | TTCATCTTCCTTTCACACATTTGAGTTGATAAATCAACTATTG       |
|            | 772   |            | 772   |
|            |   |            |   |
| Db         | 16036   | Db         | 16036   |
|            | AAATATTGCAACTTTAACAAATAAATTAATAATTAATAATTA        |            | AAATATTGCAACTTTAACAAATAAATTAATAATTAATAATTA        |
|            | 16095   |            | 16095   |
|            |   |            |   |
| QY         | 833   | QY         | 833   |
|            | TAAGAATCAAATTTCTAACTCGACTTAATGCACTTACATCA         |            | TAAGAATCAAATTTCTAACTCGACTTAATGCACTTACATCA         |
|            | 892   |            | 892   |
|            |   |            |   |
| Db         | 16156   | Db         | 16156   |
|            | ATTAAATTAATTATTATTATTTCACAAATTATCATTCAAAAT        |            | ATTAAATTAATTATTATTATTTCACAAATTATCATTCAAAAT        |
|            | 952   |            | 952   |
|            |   |            |   |
| QY         | 893   | QY         | 893   |
|            | GCACTTCACTTAAAGCTCAACAAATTATCATTCAAAAT            |            | GCACTTCACTTAAAGCTCAACAAATTATCATTCAAAAT            |
|            | 952   |            | 952   |
|            |   |            |   |
| Db         | 16216   | Db         | 16216   |
|            | TTTGTGATATTGGAAATTGATATTGAAATTGATTTGATATTG        |            | TTTGTGATATTGGAAATTGATATTGAAATTGATTTGATATTG        |
|            | 16275   |            | 16275   |
|            |   |            |   |
| QY         | 953   | QY         | 953   |
|            | TCTATAATTCTTGGATTCACAAAAGAACAGGAATTGGAA           |            | TCTATAATTCTTGGATTCACAAAAGAACAGGAATTGGAA           |
|            | 1012  |            | 1012  |
|            |   |            |   |
| Db         | 16276   | Db         | 16276   |
|            | AATTAAATTATTATTCCTTTAAATGCTT--TCATTATTAA          |            | AATTAAATTATTATTCCTTTAAATGCTT--TCATTATTAA          |
|            | 16332   |            | 16332   |
|            |   |            |   |
| QY         | 1013  | QY         | 1013  |
|            | CACATTCAATTGTAAAAGGGAGAGTTGTGAGGCTCATAA           |            | CACATTCAATTGTAAAAGGGAGAGTTGTGAGGCTCATAA           |
|            | 1072  |            | 1072  |
|            |   |            |   |
| Db         | 16333   | Db         | 16333   |
|            | TTTATTTATTTAGAAAAAAATTTTTTAAAGTTTAATGCTCA         |            | TTTATTTATTTAGAAAAAAATTTTTTAAAGTTTAATGCTCA         |
|            | 1132  |            | 1132  |
|            |   |            |   |
| QY         | 1073  | QY         | 1073  |
|            | TAAGTAATCTAAACAAACAAACCTTCTAAACCTTAATTA           |            | TAAGTAATCTAAACAAACAAACCTTCTAAACCTTAATTA           |
|            | 1132  |            | 1132  |
|            |   |            |   |
| Db         | 16393   | Db         | 16393   |
|            | AACTAAATAAAATAAAATAAAATAAAATAAAATAAAATAA          |            | AACTAAATAAAATAAAATAAAATAAAATAAAATAAAATAA          |
|            | 16452   |            | 16452   |
|            |   |            |   |
| QY         | 1133  | QY         | 1133  |
|            | ACAAACGGGATAAGTTAGTTAGCCCAAGCTGTTTTGGCC           |            | ACAAACGGGATAAGTTAGTTAGCCCAAGCTGTTTTGGCC           |
|            | 1192  |            | 1192  |
|            |   |            |   |
| Db         | 16453   | Db         | 16453   |
|            | AAAAAAATAAAATAAAATAAAATAAAATAAAATAAAATAA          |            | AAAAAAATAAAATAAAATAAAATAAAATAAAATAAAATAA          |
|            | 16512   |            | 16512   |
|            |   |            |   |
| QY         | 1193  | QY         | 1193  |
|            | TTTGTCTACTTTTGTGGACTTTAACCTGTTGCAACTTCT           |            | TTTGTCTACTTTTGTGGACTTTAACCTGTTGCAACTTCT           |
|            | 1252  |            | 1252  |
|            |   |            |   |
| Db         | 16513   | Db         | 16513   |
|            | TTTGTCTACTTTTGTGGACTTTAACCTGTTGCAACTTCT           |            | TTTGTCTACTTTTGTGGACTTTAACCTGTTGCAACTTCT           |
|            | 16567   |            | 16567   |
|            |   |            |   |
| QY         | 1253  | QY         | 1253  |
|            | TGTATCGCTTTTAACTAGCAACTCTGTAAAGTCCCTTC            |            | TGTATCGCTTTTAACTAGCAACTCTGTAAAGTCCCTTC            |
|            | 1312  |            | 1312  |
|            |   |            |   |
| Db         | 16568   | Db         | 16568   |
|            | TTAACATTTTTTCAATTATTTAACATACACATACATAA            |            | TTAACATTTTTTCAATTATTTAACATACACATACATAA            |
|            | 16627   |            | 16627   |
|            |   |            |   |
| QY         | 1313  | QY         | 1313  |
|            | ATCACTTC---TTGAAATATGTAATCGAAATAGTTTC             |            | ATCACTTC---TTGAAATATGTAATCGAAATAGTTTC             |
|            | 1367  |            | 1367  |
|            |   |            |   |
| Db         | 16628   | Db         | 16628   |
|            | CATATACACCTTTTACATATCATTTTACATTTACACAT            |            | CATATACACCTTTTACATATCATTTTACATTTACACAT            |
|            | 16687   |            | 16687   |
|            |   |            |   |
| QY         | 1368  | QY         | 1368  |
|            | AAAATAACCGGGTATTTCATAATTCCCAACTCTGT               |            | AAAATAACCGGGTATTTCATAATTCCCAACTCTGT               |
|            | 1406  |            | 1406  |
|            |   |            |   |
| Db         | 16688   | Db         | 16688   |
|            | TAATAATAATGATGATCAGCCGTAGAGACGTTT                 |            | TAATAATAATGATGATCAGCCGTAGAGACGTTT                 |
|            | 16726   |            | 16726   |
|            |   |            |   |

|           |   |
|-----------|---|
| SOURCE    | Stevor; T-complex protein 1 epsilon subunit; telomere; var.   |
| ORGANISM  | Plasmodium falciparum 3D7.  |
| REFERENCE | Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium   |
| AUTHORS   | Bowman, S., Lawson, D., Brown, D., Chillingworth, T., Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Feltwell, T., Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S., Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kytes, S., McLeod, S., Moule, S., Mungall, K., Murphy, L., Oliver, K., Quail, M.A., Rajandream, M.A., Rutter, S., Squares, R., Squares, R., Sulston, J.E., Whitehead, S., Skelton, J., Woodward, J.R., Newbold, C. and Barrell, B.G.        |
| TITLE     | The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum   |
| JOURNAL   | Nature 400 (6744), 532-538 (1999)   |
| MEDLINE   | 99376085  |
| PubMed    | 10448855  |
| REFERENCE | 2 (bases 1 to 253305)   |
| AUTHORS   | Unpublished   |
| REFERENCE | 3 (bases 1 to 253305)   |
| AUTHORS   | Lawson, D., Bowman, S. and Barrell, B.  |
| TITLE     | Direct Submission   |
| JOURNAL   | Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium   |
| COMMENT   | The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK  |
| COMMENT   | On or before May 14, 2001 this sequence version replaced 91:2982551, 91:2982541, 91:2982544, 91:2984453, 91:2894464, 91:2982556, 91:2894491, 91:2894496, 91:2982569, 91:2982563, 91:2892564, 91:2873766, 91:2894496, 91:2882577, 91:269594, 91:2894502, 91:2982538, 91:2894389, 91:2894388, 91:4725992. For more information about this sequence or the Malaria Project see <a href="http://www.sanger.ac.uk/Projects/P_falciparum/">http://www.sanger.ac.uk/Projects/P_falciparum/</a> |
| FEATURES  | Location/Qualifiers   |

```

source
organism="Plasmodium falciparum 3D7"
strain="3D7"
/db_xref="taxon:36329"
/clone="MAL3P7"
/chromosome="3"
/gene="MAL3P7.1"
/locus="PFC0860w", MAL3P7.1"
/note="PFC0860w (MAL3P7.1). Kinesin-related protein.
1200 aa: similarity: to kinesin-related protein.
D.melanogaster kinesin-like protein 67A (TRP91945)
Score: 664, sum P(2) = 1.9e-67; 28% identity in 707
overlap, predicted using hexboxn, Pfam: match to PF
kinesin, Kinesin motor domain Score 311.54"
/codon_start=1
/product="kinesin-related protein"

```

|                           |        |                  |           |                |
|---------------------------|--------|------------------|-----------|----------------|
| Query Match               | 4.78;  | Score 69.2;      | DB 3;     | Length 253305; |
| Best Local Similarity     | 42.68; | Pred. No. 0.021; |           |                |
| Matches 416; Conservative | 0;     | Mismatches 558;  | Indels 2; | G              |

Search completed: February 23, 2003, 22:57:37  
Job time : 6123 secs

|    |        |  |        |
|----|--------|--|--------|
| 2Y | 144    | AATTCTCTGTCAGGAGGATAATACAGAAAATACGAAACCTCCAAAAAAATT        | 203    |
| Db | 177318 | AATTATGGATAAAGAATTTCATATAATCAAAACATGGAGATCCTTCAAGAAAGGC    | 177377 |
| 2Y | 204    | TGGACCCUTACACGACATAAATTGGCGATAAACCTGCCATAATAAAACTTGAAACA   | 263    |
| Db | 177378 | TATAAGTAAAGTAAACATUTTGTTGATCAACTAAAGGGTTGAAATAATGAAATAA    | 177437 |
| 2Y | 264    | TACGATATGTTATCTTTCTACATCTGAAATTTCCTTTTAACATTATGAAACA       | 323    |
| Db | 177438 | GAATAATCAGCTAAUTTAATTTGAGAAAATTGATTACAAATGAGGAAATA         | 177497 |
| 2Y | 324    | ATTGAAAAAAAGGAAATGAAAAGTAAAGGTGCTAACCATGAAATAATAGCTA       | 383    |
| Db | 177498 | AGCACACAAAAAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAATAG       | 177557 |
| 2Y | 384    | ACGTTTTCTCTGTCGCTTTACTAAAGGAAATAACAAAGGAAATTAGCGATAACCTG   | 443    |
| Db | 177558 | AGGAAGAAAAAGATCATGACAAATTCTATTTGTTATATTTTTTTTTG            | 177617 |
| 2Y | 444    | CGTAAGGGTGTCAACAAAATATTTGACCGTTAGCTCTATAGAAAATATAACAACCTAA | 503    |
| Db | 177618 | TTTTTATTTATAGAATTTAGGATATTATTTGAAATTACAAAATGAA             | 177677 |
| 2Y | 504    | TCTCTTGGAAATTCTCTATATAATAGAAATCCCTCTCACAGTGAACTGAAATTTC    | 563    |
| Db | 177678 | AGGACATATATAATATAATATAATTTGTTGAGACATATAATTTGAACTTATA       | 177737 |